Barrier Analysis And Accident Prevention

ERIK HOLLNAGEL

CSELAB, Department of Computer and Information Science, University of Linköping, 58183 Linköping, Sweden, eriho@ida.liu.se

Dedale SA, 4 place de Londres, 95727 Roissy CDG Cedex, France, ehollnagel@dedale-sa.com

Introduction

Accident – and incident – analyses are usually aimed at finding the set of causes that is necessary and sufficient to explain what happened. In many cases the determination of the cause, however, reflects the interests of the stakeholders as much as what actually happened. As Perrow (1986) noted:

> "Formal accident investigations usually start with an assumption that the operator must have failed, and if this attribution can be made, that is the end of serious inquiry. Finding that faulty designs were responsible would entail enormous shutdown and retrofitting costs; finding that management was responsible would threaten those in charge, but finding that operators were responsible preserves the system, with some soporific injunctions about better training" (p. 146).

Even if the investigation into what caused the accident is more open-minded than Perrow assumes, focusing on finding the cause easily detracts interest from the other conditions that contributed to the accident. A proper accident analysis should not only look for possible specific causes, but also for the general system conditions at the time of the accident, specifically the barriers that may have failed. An understanding of the nature of barriers and defences, and a method for analysing and classifying their functions and failures is important as a basis for taking positive steps to prevent future accidents.

The Efficiency Of Barriers

Large accidents invariably represent an unlikely combination of many individual factors or causes. The prevention of a recurrence of the accident, or of similar accidents, is therefore more likely to be achieved by improving the barriers. The reasoning behind this is simple. Since serious accidents are due to coincidences among multiple factors and conditions, there are no simple "root causes". Indeed, there may be several equivalent conditions that can turn an event into an accident. Removing just one or a few of these – and moreover the ones that attracted most attention during the analysis – will not guarantee against a recurrence. The solution to eliminate causes is only appropriate if it can be assumed that a sequential description of the accident is valid, e.g. a domino type of model.

Unlike the elimination of causes, barriers are effective because they can protect against a specific type of effect regardless of why it came about. To take a very simple example, an umbrella is effective against water in the air (precipitation) regardless of whether the source is rain, sleet, a fountain, a waterfall, a sprinkler, etc. Similarly, a sprinkler system is effective against fire regardless of the origin of the fire, and a parachute can be used to save lives

regardless of the reason for needing to escape. Accident analyses should therefore not only look for causes but also try to find barriers that have failed or barriers that were missing, and in both cases go further to determine why they failed or were missing. As a response, introducing new or improved barriers is an effective way of preventing a type of accident from occurring again – provided, of course, that the barriers are effective and that they do not adversely affect the accomplishment of the task.

Accident models

Thinking about accidents involves a number of accident models, which are stereotypical ways of explaining how accidents occur. Although there are many individual models, they seem to fall into the three types summarised in **Table 1**. The simplest types of accident models describe the accident as the result of a sequence of events that occur in a specific order. The description of the sequence may either represent the scenario as a whole, or only the events that went wrong. The first type is illustrated by the so-called domino model (Heinrich, 1931), which depicts the accident as a set of dominos that tumble because of a unique initiating event. In this model the dominos that fall represent the action failures, while the dominos that remain standing represent the normal events. The outcome is a necessary consequence of one specific event, and it can therefore be considered a deterministic model. Another example is the Accident Evolution and Barrier model (Svenson, 1991), which only describes the sequence of events – or rather barriers – that failed. This puts the focus on what went wrong, but leaves out additional information that is potentially important. Sequential models need, of course, not be limited to a single sequence of events but may be represented in the form of hierarchies such as the traditional event tree and networks such as Critical Path models (Programme Evaluation and Review Technique or PERT) or Petri network.

Sequential models are attractive because they are easy to understand and represent graphically, but they suffer from being oversimplified. The epidemiological models may therefore be a better alternative. As the name implies, these models describe an accident in analogy with a disease, i.e., as the outcome of a combination of factors, some manifest and some latent, that just happen to exist together in space and time. The classical example of that is the description of latent conditions by Reason (1990). Other examples are models that consider barriers and carriers, and models of pathological system (organisation) states. Epidemiological models are valuable because they, at least, provide a basis for discussing the complexity of accidents that overcome the limitations of sequential models. Unfortunately, epidemiological models are rarely stronger than the analogy, i.e., they are difficult to specify in further detail, even though the concept of pathogens allows for a set of methods that can be used to characterise the general “health” of a system (Reason, 1997).

A third type of models is the so-called systemic model. The name denotes that these models endeavour to describe the characteristic performance on the level of the system as a whole, rather than on the level of specific cause-effect “mechanisms”. Systemic models are found in control theory, in the analogical use of the Brownian movement, in chaos models, and in coincidence models. A token for the latter type is the Swiss cheese analogy (Reason, 19xx), although this is not a model in the usual meaning of the term. In general, systemic models emphasise the need to base accidents analysis on an understanding of the functional characteristics of the system, rather than on assumptions or hypotheses about internal mechanisms as provided by standard representations of e.g. information processing or failure pathways.

Table 1: The main types of accident models.

Model type	Example
Sequential models	Linear chain of events (domino) Tree models Network models
Epidemiological models	Latent conditions Carrier-barriers Pathological systems
Systemic models	Control theoretic models "Brownian" movement models Coincidence models

Each of the types characterised above has consequences for how the accident analysis takes place, and what the outcome is.

- For the sequential models the accident analysis becomes a search for recognisable, specific causes and well-defined cause-effect links. The underlying assumption is that causes, once they have been found, can be eliminated or enclosed.
- In the case of the epidemiological models, the accident analysis becomes a search for "carriers" and latent conditions, as well as for reliable indications of general system "health". In this case the underlying assumption is that defences and barriers can be strengthened to prevent future accidents from taking place.
- Finally, for the systemic models the analysis becomes a search for unusual dependencies and common conditions that turn into coincidences. This reflects the belief that the essential variability of a system can be detected and controlled.

It is clearly not possible to analyse an accident without having some kind of underlying model. The quest should therefore not be for a "model-free" analysis, but rather for an approach where the model constraints are as light as possible. At the same time the model should be detailed enough to support an explicit and consistent method for analysis, as well as being a basis for recommending appropriate responses.

The relativity of causes

One of the fundamentals of Western thinking is causality principle, which permeates both moral philosophy and scientific thinking. In the latter domain the paradigmatic example is Newton's laws. Since we generally "know" that every cause has an effect, we automatically assume that every effect also has a cause, and furthermore that this cause can be found by deductive reasoning. According to this way of thinking an accident constitutes an effect, and it must therefore be possible to find the preceding cause – or set of causes. The common accident models as discussed above, especially the sequential models, reinforce this assumption.

A cause is, however, not an absolute condition or state of the system that is waiting to be discovered and which therefore can be determined with a high degree of certainty. As suggested by Perrow (1986), the search for a cause is often far from objective. Even if it is acceptably objective, there are always practical constraints of e.g. resources or time that limit the search. Every analysis must stop at some time, and the criterion is often determined by interests that are quite remote from the scientific purpose of the accident investigation. As Woods et al. (1994) have pointed out, a cause is always the result of a judgement made in hindsight, and therefore benefits from the common malaise of *besserwissen.* More precisely, a cause can be defined as the *ex post facto* identification of a limited set of aspects of the situation that are seen as the necessary and sufficient conditions for the effect(s) to have occurred. A "cause" usually has the following characteristics:

- It can unequivocally be associated with a system structure or a function (people, components, procedures, etc.).

- It is possible to do something to reduce or eliminate the cause within accepted limits of cost and time. This follows partly from the first characteristic, or rather, the first characteristic is a necessary condition for the second.
- The cause conforms to the current "norms" for explanations, as encapsulated by the theories that are part of the common lore. For instance, before the 1960s it was uncommon to use "human error" as a cause, while it practically became *de rigueur* during the 1970s and 1980s. Later on, in the 1990s, the notion of organisational accidents became accepted, and the norm for explanations changed once more.

The reason for making these possibly obvious points is to emphasise that the determination of the "cause" is a relative rather than an absolute process, and that it represents pragmatic rather than scientific reasoning. The outcome of an accident analysis should be treated with care, especially when it comes to thinking about responses to the accident. Knowing how systems have failed in the past is essential for predicting how they may fail in the future: design encapsulates experience. There is, however, an unfortunate disparity between accident analysis and accident prediction. In the case of analysis, it is by now commonly accepted that models should be of the epidemiological or systemic rather than of the sequential type. Yet in the case of design and prediction most models are of the sequential type. This can easily be demonstrated by referring to such widespread models as the PSA event tree, the family of fault tree or cause-consequence models, the failure mode and effect analysis methods, etc. Since this disparity severely hampers our ability to anticipate failures that may occur, there is a need to develop accident models that are able to capture the complexity of coincidences and which can be applied to prediction.

What Are "Errors"?

Causes are usually associated with actions, either directly or through a series of intermediate antecedent-consequent steps (Hollnagel, 1998). The notion of an action gone wrong or an "error" has been widespread, but as several people have pointed out it is a potentially misleading oversimplification (Hollnagel, 1993; Senders & Moray, 1991; Woods et al., 1994). This is easy to see if instead of using the two-sided distinction between correct and incorrect actions, we look at the following categories (cf. Amalberti, 1996).

- Actions where the actual outcome matches the intended outcome. Such actions are regarded as correctly performed actions, even though the outcome may have come about in other ways.
- Incorrectly performed actions where the failure is detected and corrected. This can happen while the action is being carried out, e.g. mistakes in typing or data entry, or immediately after, as long as the system makes a recovery possible. In these cases the actual and intended outcomes may still match, and the action is therefore often considered as correct.
- Incorrectly performed actions where the failure is detected but not corrected or recovered. Recovery can be impossible for several reasons, for instance that the system has entered an irreversible state, that there is insufficient time or resources, etc. In these cases the actual and intended outcomes do not match, and the action may therefore be characterised as an error.
- Incorrectly performed actions where the failure is detected but ignored. This may happen because the expected consequences of the failure are seen as unimportant in an absolute or relative sense. This assessment may be correct or incorrect, depending, among other things, on the users' knowledge of the system in question. If it turns out that the consequences were not negligible, the action may in retrospect be classified as an error.
- Incorrectly performed actions, which are not detected at the time, and therefore not recovered. These will as a rule lead to unwanted consequences, hence be classified as errors.

This description of five categories of action is clearly preferable to the two-way distinction between correct actions and errors. Furthermore, it is not necessarily bad if an

action is incorrectly performed. So long as the outcome does not lead to a serious and irreversible condition, the incorrectly performed action provides an important opportunity to learn. Learning cannot take place if everything is done correctly, and if there are no unexpected outcomes. (Of course, one could argue that if everything is done correctly, then there is no need to learn either.)

The extended classification of actions is consistent with the position that it is not the action in itself but rather the outcome that is incorrect. (Indeed, the verdict of an incorrectly performed action is clearly a relative rather than an absolute judgement.) It must be acknowledged that human performance individually and in groups, as well as the performance of technological artefacts, always is variable. Sometimes the variability remains within acceptable limits, but at other times it becomes so large that it leads to unexpected and unwanted consequences. In both cases, however, the basis for the performance variability is the same, and that which makes us classify one action as an "error" and the other as not are the outcomes. It follows from this view that rather than trying to identify specific causes and eliminating them, we should try to detect the performance variability in order to control it, either by reducing it at source or by protecting against the outcomes. In both cases managing the performance variability becomes more important than searching for and eradicating errors.

Barrier Systems And Barrier Functions

In relation to accidents, a barrier is an obstacle, an obstruction, or a hindrance that may either (1) prevent an action from being carried out or an event from taking place, or (2) prevent or lessen the impact of the consequences, for instance by slowing down uncontrolled releases of matter and energy.

In the practical work with barriers it is useful to make a distinction between barrier functions and barrier systems. A **barrier function** can be defined as the specific manner by which the barrier system achieves its purpose. Similarly, a **barrier system** can be defined as the basis for the barrier function, i.e., the characteristics of a system without which the barrier function could not be accomplished. It is possible from this basis to develop a systematic description of various types of barrier systems and barrier functions, which can be used as a starting point for practical methods (Hollnagel, 1999).

Despite the importance of the barrier concept, the accident literature only contains a small number of studies (Leveson, 1995; Svenson, 1991 & 1997; Taylor, 1998 and Trost & Nertney, 1985). Other uses has been in relation to the notion of defences, for instance by Reason (19xx) and in the Japanese approach to accident prevention named *hiyari-hatto*. The classifications proposed by these studies have been quite diverse, partly because of the lack of a common conceptual background, and partly because they have been developed for specific purposes within quite diverse fields. The most explicit attempt of developing a theory of barriers has been the work of Svenson (1991), which also was the basis for the field studies of Kecklund et al (1996).

Types Of Barrier Systems

An analytical description of barrier systems can be based on different concepts, such as their origin (e.g. whether they are created by organisations or individuals), their purpose, their location or focus (relative to e.g. the source or target), and their nature. Of these only the concept of the nature of the barrier system is rich enough to support an extensive classification. The nature of a barrier system is furthermore independent of its origin, its purpose (e.g., as preventive or protective), and its location. Although there initially may seem to be many different types of barrier systems, ranging from physical hindrances (walls, cages) to ethereal rules and laws, experience shows that the following four categories are sufficient.

- **Material barrier systems** physically prevent an action from being carried out or the consequences from spreading. Examples are buildings, walls, fences, railings, bars, cages, gates, etc. A material barrier system presents an actual physical hindrance for the action

or event in question and although it may not prevent it under all circumstances, it will at least slow it down or delay it. Furthermore, a material barrier system does not have to be perceived or interpreted by the acting agent in order to serve its purpose. A wall will prevent movement of an agent (or a substance) from one location to another even if the agent cannot see the wall – provided, of course, that it is strong enough.

- **Functional (active or dynamic) barrier systems** work by impeding the action to be carried out, for instance by establishing logical or temporal interlocks. A functional barrier system effectively sets up one or more pre-conditions that must be met before something can happen. These pre-conditions need not be interpreted by a human, but may be interrogated or sensed by technological artefacts. Functional barrier systems may not always be visible or discernible, although their presence often is indicated to human users in one way or another and may require one or more actions to be overcome.
- **Symbolic barrier systems** require an act of interpretation to achieve their purpose, hence an "intelligent" agent that can react or respond to the barrier system. Whereas a functional barrier system works by establishing a pre-condition that must be met by the acting agent or user before further actions can be carried out, limitations or constraints indicated by a symbolic barrier system may be disregarded or neglected. For instance, the railing along a road constitutes a material and a symbolic barrier system at the same time, while reflective posts or markers are only a symbolic barrier system. The reflective markers indicate where the edge of the road is but are by themselves insufficient to prevent a car from going off the road. Although all kinds of signs and signals are symbolic barriers systems, visual and auditory signals play a special role in normal work environments as part of warnings (texts, symbols, sounds), interface layout, information presented on the interface, visual demarcations, etc.
- **Immaterial** (or nonmaterial) **barrier systems** are not physically present or represented in the situation, but depend on the knowledge of the user to achieve their purpose. Immaterial barrier systems usually also have a physical representation, such as a book or a memorandum, but this is normally not present when their use is mandated. Typical immaterial barrier systems are rules, guidelines, restrictions, and laws.

The four types of barrier systems may, at a first glance, seem to be incomplete since common types such as organisational barriers and technical barriers are missing. This lack is, however, only an apparent lack since a combination of barrier systems and barrier functions will suffice to account for every type of barrier. The proposed definitions also mean that more than one barrier system may be present in the same physical artefact or object. For instance, a door may have on it a written warning and include a lock that requires a key to be opened. Here the door is a material barrier system, the written warning is a symbolic barrier system, and the lock requiring a key is a functional barrier system. It is probably the rule rather than the exception that more than one barrier system is used at the same time, at least for the first three categories.

Types Of Barrier Functions

Whereas it was possible to make do with only four different barrier systems, there are several more barrier functions. As a start, barrier functions can either prevent an accident from taking place or protect against the consequences. The overall functions of prevention and protection can be further specialised, depending on the domain and on the type of barrier system. All four types of barrier system, for instance, can accomplish prevention – although not with the same degree of efficiency. Protection, on the other hand, cannot be provided by symbolic or immaterial barrier systems.

In the development of an accident, prevention refers to what may be done before the accident occurs while protection refers to what may be done after. Considering the following four stages can refine this distinction:

- Steady-state performance. Here the main concern is to monitor how the system performs. While monitoring does not constitute a class of barrier function as such,

effective monitoring can prevent accidents. Monitoring may involve functions such as observation, confirmation, managing, and recording.

- Pre-accident build-up. At this stage the main objective is to detect variations or deviations in system performance. This is a genuine part of accident prevention, and many system design features (technological and organisational) relate to that. Detection may involve functions such as identification, verification, and questioning or probing, and typically relies heavily on functional barrier systems.
- When the accident happens, the first stage of protection is to reduce or deflect the immediate consequences. Some characteristic functions here are attenuating, partitioning, and reducing the direct effects, as well as strengthening defences and resources.
- The post-accident or recovery period constitutes the second stage of protection. This covers a longer time period and can be seen as a way of correcting what went wrong, involving replacement, modification and improvement of both barrier systems and barrier functions.

It is possible to develop this classification further, and thereby provide the basis for a comprehensive approach to barrier analysis and accident prevention. The principles for this are outlined in **Table 2**, but space prohibits a full treatment here. The high-level barrier functions of monitoring – detecting – deflecting – correcting, describes the characteristics stages in the development of an accident. Each of the high-level barrier functions can be implemented by one or more barrier systems; the choice of the most effective solution depends on the requirements from the stakeholders as well as the prerequisites for the individual functions. A correction that is implemented by an immaterial barrier system, such as a new rule, can be effectuated very quickly and inexpensively. On the other hand, immaterial barriers are rarely effective in the long run, since they require a high degree of compliance by those involved.

Table 2: Barrier functions and accident prevention.

	Barrier system			
Barrier function (high level)	**Material**	**Functional**	**Symbolic**	**Immaterial**
Monitoring	Monitoring involves an act of reasoning, hence a function. It cannot be provided by a material barrier system .	Monitoring can be provided by a functional barrier system, which checks (logical) conditions	Monitoring can be provided by a symbolic barrier system, which interprets signs and indicators	Monitoring involves an act of reasoning, hence a function. It cannot be provided by an immaterial barrier system .
Detecting	Detection cannot be provided by a material barrier system, since it involves an act of reasoning, hence a function.	Detection is usually functional, and can be implemented by technology with or without human collaboration.	Detection by humans (unaided or aided by machines) can be provided by symbols and interpretation.	Detection cannot be provided by an immaterial barrier system, since it requires an act of identification / interpretation
Deflecting or reducing	Deflection is usually provided by a material barrier system, such as a firewall, a fire belt, etc.	Deflection can be provided by a functional barrier system, such as interlocks, airbags or sprinklers.	Deflection cannot be provided by a symbolic barrier system, since it means changes in the direction of matter and energy	Deflection cannot be provided by an immaterial barrier system, since means changes in the direction of matter and energy
Correcting	Correction can be provided by a material barrier system, such as restoring.	Correction can be provided by a functional barrier system, such as developing new interlock	Correction can be provided by a symbolic barrier system, such as developing new signs and symbols	Correction can be provided by an immaterial barrier system, such as instituting new laws

From "Error" Management To Performance Variability

The argument of this paper is that it may be more efficient to prevent accidents through the judicious use of barrier systems and barrier functions than to identify and eliminate specific causes. This corresponds to the view that causes represent an *ex post facto* attribution, based on an oversimplified understanding of the nature of accidents. Given that accidents more often are due to complex coincidences than well-defined cause-effect relations, it makes sense to approach accident prevention by managing the variability of the system. This in turn can be accomplished by considering the various characteristic stages of an accident, as described above, and by applying the barrier functions that are most effective at each stage. Many of the functions necessary for managing system performance variability are already in place in both the technological and organisational area. They have, however, usually been implemented in a piecemeal fashion and without the benefit of an overall perspective. The further application of performance variability management is therefore more a question of using existing principles in an integrated fashion than of inventing completely new ways of doing things. Efforts in this direction are currently underway in a number of domains, ranging from industrial manufacturing and production to traffic management.

References

Amalberti, R. (1996). *La conduite des systèmes à risques*. Paris: PUF.

Heinrich, H. (1931). *Industrial accident prevention*. New York: McGraw-Hill.

Hollnagel, E. (1998). *Cognitive reliability and error analysis method – CREAM*. Oxford: Elsevier Science.

Hollnagel, E. (1993). *Human reliability analysis: Context and control*. London: Academic Press.

Hollnagel, E. (1999). Accidents and barriers. In J.-M. Hoc, P. Millot, E. Hollnagel & P. C. Cacciabue (Eds.), Proceedings of. *Lez Valenciennes, 28*, 175-182. (Presses Universitaires de Valenciennes.)

Kecklund, L. J., Edland, A, Wedin, P. & Svenson, O. (1996). Safety barrier function analysis in a process industry: A nuclear power application. *Industrial Ergonomics, 17*, 275-284.

Leveson, N. (1995). *Safeware. System safety and computers*. Reading, MA: Addison-Wesley Publishing Company.

Perrow, C. (1986, 3rd ed). *Complex organizations: A critical essay*. New York: Random House.

Reason, J. (1990). The contribution of latent human failures to the break down of complex systems. *Philosophical Transactions of the Royal Society (London)*, Series B. 327: 475-484.

Reason, J. T. (1997). *Managing the risks of organizational accidents*. Aldershot, UK: Ashgate.

Senders, J. W. & Moray, N. P. (1991). *Human error. Cause, prediction, and reduction*. Hillsdale, NJ.: Lawrence Erlbaum.

Svenson, O. (1991). The accident evolution and barrier function (AEB) model applied to incident analysis in the processing industries. *Risk Analysis, 11*(3), 499-507.

Svenson, O. (1997). Safety barrier function analysis for evaluation of new systems in a process industry: How can expert judgment be used? In: *Proceedings of Society for Risk Analysis Europe Conference*, Stockholm, June 15-18, 1997.

Taylor, R. J. (1988). *Analysemetoder til vurdering af våbensikkerhed*. Glumsø, DK: Institute for Technical Systems Analysis.

Trost, W. A. & Nertney, R. J. (1985). *Barrier analysis* (DOE 76-45/29). Idaho Falls, Idaho: EG&G Idaho, Inc.

Woods, D. D., Johannesen, L. J., Cook, R. I. & Sarter, N. B. (1994). *Behind human error: Cognitive systems, computers and hindsight*. Columbus, Ohio: CSERIAC.